# The structure of optimal solutions to the submodular function minimization problem

Collette Coullard [1]

June 23, 2003

**Abstract**. In this paper, we study the structure of optimal solutions to the submodular function minimization problem. We introduce *prime sets* and *pseudo-prime sets* as basic building block of minimizer sets, and investigate composition, decomposition, recognition, and certification of prime sets. We show how Schrijver's submodular function minimization algorithm can be modified to construct in polynomial time a prime or pseudoprime decomposition of the ground set. We also show that the final vector $x$ obtained by this algorithm is an extreme point of the polyhedron $P := \{x \in \mathbf{R}^V : x \le 0;\ x(A) \le f(A) \text{ for all } A \subseteq V\}$.

## 1  Introduction

The problem of minimizing a submodular function has been known to be polynomially solvable since 1981, when Grötschel, Lovász, and Schrijver [10] showed it could be solved via the ellipsoid method. Until 1999, a purely combinatorial polynomial algorithm for this problem was unknown. In 1999, Schrijver [18] and independently Iwata, Fleischer, and Fugishige [14] published such algorithms. The paper by McCormick [16] gives an excellent account of the history and recent developments of submodular function minimization.

The known certificate of optimality for submodular function minimization is to exibit a vector $x$, say, as a convex combination of a set of vectors each obtained via the greedy algorithm applied using some chosen ordering of the ground set. A set $U$ is shown to be a minimizer if all of the negative and none of the positive components of $x$ are members of $U$, and $U$ is a lower ideal in each of the chosen orderings.

This *convex certificate of optimality* is polynomially verifiable, since the collection of chosen orderings can be taken to be of size no greater than the ground set. Thus, the certificate is verifiable using a quadratic number of function calls and elementary operations.

The motivation for the present work was to find a different certificate of optimality, one that doesn't involve convex combinations, and one, hopefully, that could be verified in linear time.

While this paper doesn't succeed in this goal, it does provide some new insights about the structure of optimal solutions. First, we introduce two slightly different basic building blocks for minimizer sets, which we call *prime sets* and *pseudo-prime sets*. We show in Theorem 12 that every minimizer set decomposes into a sequence of prime sets, each of which adds nonnegative marginal value (with respect to its predecessors in the sequence), and we also show in Theorem 19 it decomposes similarly into a sequence of pseudo-prime sets.

| | | Form Approved OMB No. 0704-0188 |
|---|---|---|

# Report Documentation Page

| 1. REPORT DATE **JUN 2003** | 2. REPORT TYPE | 3. DATES COVERED **00-00-2003 to 00-00-2003** |
|---|---|---|
| 4. TITLE AND SUBTITLE **The structure of optimal solutions to the submodular function minimization problem** | | 5a. CONTRACT NUMBER |
| | | 5b. GRANT NUMBER |
| | | 5c. PROGRAM ELEMENT NUMBER |
| 6. AUTHOR(S) | | 5d. PROJECT NUMBER |
| | | 5e. TASK NUMBER |
| | | 5f. WORK UNIT NUMBER |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Northwestern University,Department of Industrial Engineering and Management Sciences,Evanston,IL,60208** | | 8. PERFORMING ORGANIZATION REPORT NUMBER |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | 10. SPONSOR/MONITOR'S ACRONYM(S) |
| | | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) |

12. DISTRIBUTION/AVAILABILITY STATEMENT
**Approved for public release; distribution unlimited**

13. SUPPLEMENTARY NOTES

14. ABSTRACT
**see report**

15. SUBJECT TERMS

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT **unclassified** | b. ABSTRACT **unclassified** | c. THIS PAGE **unclassified** | **Same as Report (SAR)** | **24** | |

**Standard Form 298 (Rev. 8-98)**
Prescribed by ANSI Std Z39-18

Next we go on to investigate further prime sets and pseudo-prime sets, with the goal of finding a certificate for recognizing them. If we could find such a certificate, then this would give a certificate for minimizers.

The motivating idea can be illustrated on a 2-element example. Consider the function $f$ on ground set including subset $\{a, b\}$, and having $f(\{a\}) = 1$, $f(\{b\}) = 1$, and $f(\{a, b\}) = -1$. The set $A$ is contained in every minimizer of $f$. $A$ is an example of both a prime set and a pseudoprime set. This example can be extended to one in which element $b$ is replaced by a subset $B$ with the following properties: $f(B) = 1$; $f(B \cup \{a\}) = -1$; and $f(C \cup \{a\}) > f(\{a\}) = 1$, for every proper subset $C \subset B$. Again, we can conclude that $A = B \cup \{a\}$ is contained in every minimizer of $f$.

The previous paragraph illustrates a composition of ordered prime sets in both cases: In the first case, $\{a\}$, with positive marginal value (with respect to $\emptyset$), and $\{b\}$, with negative marginal value (with respect to $\{a\}$) compose into $A$, which itself has negative marginal value; in the second case, $\{a\}$, with positive marginal value and $B$, with negative marginal value (with respect to $\{a\}$) compose into $A$, with negative marginal value.

We show in Theorem 16 that every prime set with negative marginal value decomposes into such a sequence of prime sets, where the first member has positive marginal value and the rest have non-positive marginal value. In the analog for pseudoprime sets, Theorem 20, we have that every pseudoprime set with negative marginal value decomposes into a sequence of exactly 2 pseudoprime sets, the first having non-positive marginal value, and the second having non-negative marginal value with respect to the first.

These 2 theorems give some hope that we could recognize prime sets or pseudoprime sets. The final step would be a composition theorem that gives a verifiable list of conditions under which a sequence of 2 prime (or pseudoprime) sets composes into a single prime (or pseudoprime) set.

To that end, we offer Theorems 17 and 21. Both these theorems have set $A$ with $f(A) < 0$ partitioned into a sequence $\{A_1, A_2\}$, of prime (pseudoprime) sets, each of which has nonnegative marginal value with respect to the empty set and non-positive marginal value with respect to the other. In each case we can conclude that $A$ is contained in every minimizer of $f$, but we cannot conclude that $A$ is prime (pseudoprime).

We do have a certificate to verify that a set $A$ is pseudoprime. Unfortunately, this certificate, stated in Theorem 25, does involve forming $|A|$ convex combinations, and can therefore take $O(|A|^3)$ function calls and elementary operations.

The remainder of the paper deals with the algorithmic question of finding a pseudoprime decomposition. We show that Schrijver's algorithm for submodular function minimization can be extended to produce such a decomposition. We also show that when it finishes, the components of the decomposition are exactly the strongly connected components of the associated auxiliary digraph.

Finally, we show that the vector $x$ produced by this algorithm provides an extreme point optimal solution to an associated linear programming problem.

# 2 Definitions

In this section we present necessary definitions.

Function $f$ defined on subsets of finite set $V$ is *submodular* if, for every pair of subsets $A$ and $B$ of $V$, the following inequality is satisfied:

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B). \tag{1}$$

Equivalently, for every pair of subsets $A$ and $B$ of $V$, with $A \subseteq B$, and every element $v \in V$, the following inequality is satisfied:

$$f(A + v) - f(A) \geq f(B + v) - f(B), \tag{2}$$

where, here and throughout this paper, the expressions $A \cup \{v\}$ and $A - \{v\}$ are abbreviated by $A + v$ and $A - v$, respectively, for $A \subseteq V$ and $v \in V$.

Given $A \subseteq V$ and $v \in V - A$, the *marginal value of $v$ with respect to $A$*, denoted $m(v, A)$, is $f(A + v) - f(A)$. Thus, condition (2) can be phrased as follows: The marginal value of any element is non-increasing on any sequence of sets that is ordered by inclusion. The *marginal value* of subset $B \subseteq V$ with respect to $A$ is $m(B, A) := f(A \cup B) - f(B)$. The marginal value of any subset is similarly non-increasing.

The *submodular function minimization problem* is to find $S^* \subseteq V$ with $f(S^*) \leq f(S)$, for every $S \subseteq V$. Given sets $S, A \subseteq V$, we will say $S$ is a *minimizer for $f$ on $A$* if $S \subseteq A$ and $f(S) \leq f(B)$, for every $B \subseteq A$. If no set $A$ is specified, then it means $S$ is a minimizer for $f$ on $V$.

We assume throughout that $f$ is normalized (by subtracting $f(\emptyset)$ from each $f(S)$, if necessary), so that $f(\emptyset) = 0$. This normalization maintains submodularity and does not affect the optimal solution to the submodular function minimization problem.

Given a submodular function $f$ on $V$, and given $S \subseteq V$, the submodular function obtained by *contracting $S$*, denoted $f_{V/S}$, is defined as follows:

$$f_{V/S}(P) := f(P \cup S) - f(S), \text{ for each } P \subseteq V - S.$$

Note that $f_{V/S}(\emptyset) = 0$. Let $m_{V/S}$ denote the marginal value function associated with $f_{V/S}$.

Assume $\{A, B, C, D\}$ is a partition of $V$, and Then, directly from definitions, we have:

$$
\begin{aligned}
m_{V/A}(C, B) &= f_{V/A}(C \cup B) - f_{V/A}(B) \\
&= f(C \cup B \cup A) - f(A) - f(B \cup A) + f(A) \\
&= m(C, B \cup A). \tag{3}
\end{aligned}
$$

Let $\prec$ be a total order of $V$. For $v \in V$, the *closed lower ideal* defined by $v$, denoted $[v]^\prec$, is the set $\{u \in V \ : \ u \preceq v\}$, and the *open lower ideal* defined by $v$, denoted $[v)^\prec$, is the set $\{u \in V \ : \ u \prec v\}$.

Suppose $V$ is partitioned into a sequence of sets:

$$S_1, S_2, \ldots, S_k, T_1, T_2, \ldots, T_r, \tag{4}$$

and we have a total order $\prec$ on $V$. Then $\prec$ is *compatible* with the sequence (4) if for every $u, v \in V$, if $u \in S_i$ and $v \in S_{i+1} \cup \cdots \cup T_r$, then $u \prec v$; and if $u \in T_i$ and $v \in T_{i+1} \cup \cdots \cup T_r$, then $u \prec v$. Define

$$[T_i)^\prec := S_1 \cup \ldots \cup S_k \cup T_1 \cup \ldots \cup T_{i-1}.$$

When no total order $\prec$ is specified, we will omit the superscript and just write $[S_i)$ or $[T_i)$.

## 2.1 Functions $f$ and $\bar{f}$

Assume function $f$ on subsets of $V$ is submodular and $f(\emptyset) = 0$. Define function $\bar{f}$ on subsets $A$ of $V$ as follows:

$$\bar{f}(A) := f(V - A) - f(V).$$

It is straightforward to see that $\bar{f}$ is submodular and $\bar{f}(\emptyset) = 0$. Moreover,

$$S^* \text{ is a minimizer of } f \text{ on } V \Leftrightarrow V - S^* \text{ is a minimizer of } \bar{f} \text{ on } V. \tag{5}$$

Let $\bar{m}$ denote the marginal value function defined by $\bar{f}$. That is, for each $B \in V$ and $A \subseteq V - B$,

$$\bar{m}(B, A) := \bar{f}(A \cup B) - \bar{f}(A).$$

The following relationships follow directly from the definitions:

Let $\{B, C, D\}$ be a partition of $V$.

$$
\begin{aligned}
m(C, B) &= f(B + C) - f(B) \\
&= \bar{f}(D) + f(V) - \bar{f}(C + D) - f(V) \\
&= -\bar{m}(C, D)
\end{aligned}
\tag{6}
$$

Let $\{A, B, C, D, E\}$ be a partition of $V$.

$$
\begin{aligned}
m_{V/A}(C, B) &= f(B \cup C \cup A) - f(A) - f(B \cup A) + f(A) \\
&= \bar{f}(D \cup E) + f(V) - \bar{f}(C \cup D \cup E) + f(V) \\
&= \bar{f}(D \cup E) - \bar{f}(C \cup D \cup E) \\
&= \bar{f}(D \cup E) - \bar{f}(E) - \bar{f}(C \cup D \cup E) + \bar{f}(E) \\
&= -\bar{m}_{V/E}(C, D)
\end{aligned}
\tag{7}
$$

## 2.2 Prime sets

Given a normalized submodular function $f$ on finite set $V$, a subset $\emptyset \neq A \subseteq V$ is called *prime*, with respect to $f$, if $f(B) > 0$ and $m(B, A - B) < 0$, for every $\emptyset \neq B \subset A$.

**Lemma 1** *Let $\{C, A, D\}$ be a partition of $V$ in which $A \neq \emptyset$. $A$ is prime with respect to $f_{V/C}$ if and only if $A$ is prime with respect to $\bar{f}_{V/D}$.*

PROOF Let $\emptyset \neq B \subset A$. By (7), we have that

$$
\begin{aligned}
f_{V/C}(B) &= m_{V/C}(B, \emptyset) \\
&= -\bar{m}_{V/D}(B, A - B),
\end{aligned}
\tag{8}
$$

4

and

$$m_{V/C}(B, A - B) = -\bar{m}_{V/D}(B, \emptyset)$$
$$= -\bar{f}_{V/D}(B). \tag{9}$$

Therefore,

$$f_{V/C}(B) > 0 \iff \bar{m}_{V/D}(B, A - B) < 0,$$

and

$$m_{V/C}(B, A - B) < 0 \iff \bar{f}_{V/D}(B) > 0.$$

Together, these give $A$ is prime with respect to $f_{V/C}$ if and only if $A$ is prime with respect to $\bar{f}_{V/D}$, as desired. □

**Lemma 2** *Let $\{C, A, D\}$ be a partition of $V$ in which $A \neq \emptyset$. If $f_{V/C}(A) \leq 0$, then $A$ is prime with respect to $f_{V/C}$ if and only if $f_{V/C}(B) > 0$, for every $\emptyset \neq B \subset A$. If $f_{V/C}(A) \geq 0$, then $A$ is prime with respect to $f_{V/C}$ if and only if $m_{V/C}(B, A - B) < 0$, for every $\emptyset \neq B \subset A$.*

PROOF We prove the first statement. The second follows from the first, using Lemma 1 and the relationship between $f$ and $\bar{f}$. The "only if" direction follows immediately from the definition of prime. For the "if" direction, let $\emptyset \neq B \subset A$. Then:

$$m_{V/C}(B, A - B) = f_{V/C}(A) - f_{V/C}(A - B) \tag{10}$$
$$< 0, \tag{11}$$

where (10) follows from the definition of marginal value, and (11) follows since $f_{V/C}(A) \leq 0$ and $f_{V/C}(A - B) > 0$. □

**Lemma 3** *Let $\{C, A, D\}$ be a partition of $V$ in which $A \neq \emptyset$ is prime with respect to $f_{V/C}$, and $f_{V/C}(A) \leq 0$. If $f(A) \leq 0$, then $A$ is prime with respect to $f$.*

PROOF By Lemma 6, it suffices to show that $f(B) > 0$, for each $\emptyset \neq B \subset A$. Let $\emptyset \neq B \subset A$. Then

$$f(B) = m(B, \emptyset) \tag{12}$$
$$\geq m(B, C) \tag{13}$$
$$> 0, \tag{14}$$

where (12) follows by definition of marginal value and the fact that $f(\emptyset) = 0$; (13) follows by submodularity; and (14) follows from the fact that $A$ is prime with respect to $f_{V/C}$. □

# 3   Self Minimizers

A subset $S \subseteq V$ is a *self minimizer* of $f$ if $f(S) \leq f(A)$, for every $A \subset S$. Clearly, $\emptyset$ is a self minimizer, and any minimizer of $f$ on $V$ is a self minimizer. Set $A$ is a *minimal self minimizer of $f$* if $A \neq \emptyset$, $A$ is a self minimizer of $f$, and $A$ contains no proper subset satisfying these two properties.

## 3.1 Characterizations of self minimizers and minimal self minimizers

**Lemma 4** *Let $\emptyset \neq A \subseteq V$. Then $A$ is a self minimizer of $f$ if and only if $m(B, A - B) \leq 0$, for each $\emptyset \neq B \subseteq A$.*

PROOF. For any $\emptyset \neq B \subseteq A$,

$$m(B, A - B) \;=\; f(A) - f(A - B).$$

Therefore, $m(B, A - B) \leq 0$ for all such $B$, if and only if $f(A) \leq f(A - B)$, for all such $B$, which is the definition of $A$ being a self minimizer. $\square$

**Lemma 5** *Let $\emptyset \neq B \subseteq V$. If $f(B) \leq 0$, then $B$ contains a nonempty self minimizer. In particular, any $S \subseteq B$ that is minimal with respect to both being nonempty and having $f(S) \leq 0$ is a minimal self minimizer.*

PROOF. Take $S$ to be a minimal nonempty subset of $B$ having $f(S) \leq 0$. Then since $f(S') > 0 \geq f(S)$, for every $\emptyset \neq S' \subset S$, we have that $S$ is a minimizer of $f$ on $S$, as desired. $\square$

**Lemma 6** *Let $\emptyset \neq A \subset V$. Then $A$ is a a minimal self minimizer of $f$ if and only if the following conditions hold:*

1. *$f(A) \leq 0$.*

2. *$f(B) > 0$, for every $\emptyset \neq B \subset A$.*

PROOF. ($\Leftarrow$) Let $\emptyset \neq B \subset A$. By item 2, $f(\emptyset) = 0 < f(B)$, and by item 1, $f(A) \leq f(\emptyset)$. Combining, we have $A$ is a self minimizer of $f$ and $B$ is not a self minimizer of $f$. Therefore, $A$ is a minimal self minimizer of $f$.

($\Rightarrow$) Assume $A$ is a a minimal nonempty self minimizer of $f$ and let $\emptyset \neq B \subset A$. If $f(B) \leq 0$, then by Lemma 5, $B$ contains a nonempty self minimizer, a contradiction. Therefore, item 2 must hold. Now, since $A$ is a self minimizer, $f(A) \leq f(\emptyset) = 0$ item 1 holds, as desired. $\square$

**Lemma 7** *Let $\{C, A, D\}$ be a partition of $V$ in which $A \neq \emptyset$. $A$ is a minimial self minimizer of $f_{V/C}$ if and only if $A$ is prime with respect to $f_{V/C}$, and $f_{V/C}(A) \leq 0$.*

PROOF ($\Rightarrow$) Assume $A$ is a minimal self minimizer, and let $\emptyset \neq B \subset A$. By Lemma 6, $f_{V/C}(A) \leq 0$, $f_{V/C}(B) > 0$, and $f_{V/C}(A - B) > 0$. It follows that

$$m_{V/C}(B, A - B) = f_{V/C}(A) - f_{V/C}(A - B) < 0,$$

as desired.

($\Leftarrow$) Assume $A$ is prime and $f_{V/C}(A) \leq 0$. By definition of prime, $f_{V/C}(B) > 0$, for each $\emptyset \neq B \subset A$. Now by Lemma 6, $A$ is a minimal self minimizer, as desired. $\square$

**Lemma 8** *Let $\{C, A, D\}$ be a partition of $V$ in which $A \neq \emptyset$. $A$ is a minimial self minimizer of $\bar{f}_{V/D}$ if and only if $A$ is prime with respect to $f_{V/C}$, and $f_{V/C}(A) \geq 0$.*

PROOF This follows immediately from Lemmas 1 and 8, using the fact that

$$f_{V/C}(A) \geq 0 \iff \bar{f}_{V/D}(A) \leq 0.$$

$\square$


## 3.2 Decomposition of minimizers into minimal self minimizers

**Lemma 9** *Assume $A \subseteq V$ is a self minimizer of $f$, and $C \subset A$. Then $A - C$ is a self minimizer of $f_{V/C}$.*

PROOF. Let $\emptyset \neq B \subseteq A - C$. Then

$$
\begin{align}
m_{V/C}(B, A - C - B) &= m(B, A - B) \tag{15}\\
&\leq 0, \tag{16}
\end{align}
$$

where (15) follows from (3), and (16) follows from Lemma 4 and the fact that $A$ is a self minimizer of $f$. Now, again by Lemma 4, $A - C$ is a self minimizer of $f_{V/C}$, as desired. $\square$


**Lemma 10** *Assume $A \subseteq V$. If $A$ is a self minimizer of $f$, then $A$ is contained in some minimizer of $f$.*

PROOF. Let $T$ be a minimizer of $f$. We have:

$$f(A \cup T) \leq f(A) + f(T) - f(A \cap T) \leq f(T),$$

where the first inequality holds because $f$ is submodular, and the second inequality holds because $A$ is a self minimizer of $f$. $\square$


**Lemma 11** *Assume $B \subset S \subseteq V$, that $B$ is a self minimizer of $f$, and that $S - B$ is a self minimizer of $f_{N/B}$. Then $S$ is a self minimizer of $f$.*

PROOF. Let $S' \subset S$. Then

$$
\begin{align}
f(S') &\geq f(B \cup S') + f(S' \cap B) - f(B) \tag{17}\\
&= f(S' \cap B) + f_{N/B}(S' - B) \tag{18}\\
&\geq f(B) + f_{N/B}(S - B) \tag{19}\\
&= f(S), \tag{20}
\end{align}
$$

where 17 follows by submodularity of $f$, 18 follows by definition of contraction, 19 follows because $B$ is a self minimizer of $f$ and $S - B$ is a self minimizer of $f_{N/B}$, and 20 follows by the definition of contraction. Thus, $S$ is a self minimizer of $f$, as desired. $\square$

**Theorem 12** *Assume $S$ is a minimizer of $f$. Then $S$ may be partitioned into a sequence of non-empty subsets $S_1, S_2, \ldots, S_k$, such that for each $i \in \{1, 2, \ldots, k\}$, we have that $S_i$ is a minimal self minimizer of $f_{V/[S_i)}$.*

PROOF. The proof is by induction on $|S|$. The theorem is certainly true if $|S| = 0$. Assume $|S| > 0$ and that the theorem is true for all normalized submodular functions and minimizers with fewer elements. Since $S$ is a minimizer of $f$, we know that $S$ is a self minimizer of $f$. Therefore, $S$ contains a minimal nonempty self minimizer, $S_1$, say, of $f$. By Lemma 9, $S - S_1$ is a self minimizer of $f_{V/S_1}$. By the induction hypothesis, $S - S_1$ may be partitioned into a sequence of non-empty subsets $S_2, \ldots, S_k$, such that for each $i \in \{2, \ldots, k\}$, we have that $S_i$ is a minimal nonempty self minimizer of $f_{V/(S_i]}$. This completes the proof. $\square$

**Corollary 13** *Let $f$ be a submodular function on $V$. Then there is a partition of $V$ into a sequence $S_1, S_2, \ldots, S_k, T_1, T_2, \ldots, T_l$, such that*

- *$m(S_i, [S_i)) \leq 0$ and $S_i$ is prime with respect to $f_{V/[S_i)}$, for each $i \in 1, \ldots, k$; and*

- *$m(T_i, [T_i) \geq 0$ and $T_i$ is prime with respect to $f_{V/[T_i)}$, for each $i \in 1, \ldots, l$.*

PROOF. This follows directly from Theorem 12, together with (5) and Lemma 8. $\square$

**Theorem 14** *Let $S_1, S_2, \ldots, S_k$ be a sequence of disjoint nonempty subsets of $V$ such that for each $i \in \{1, 2, \ldots, k\}$, we have that $S_i$ is a self minimizer of $f_{V/[S_i)}$. Then $S := S_1 \cup \cdots \cup S_k$ is a self minimizer of $f$. Moreover, if $S_1, S_2, \ldots, S_k$ is a maximal such sequence, then $S$ is a minimizer of $f$.*

PROOF. We begin with the proof that $S$ is a self minimizer of $f$. The proof is by induction on $k$. If $k = 1$, then the result is immediate. Assume true for all smaller $k$. Then, by the induction hypotheses, we can conclude that $S_1 \cup \cdots \cup S_{k-1}$ is a self minimizer of $f$. By Lemma 11, $S$ is a self minimizer of $f$.

Now assume $S_1, S_2, \ldots, S_k$ is a maximal such sequence; that is, $f_{V/S}$ has no nonempty self minimizer. Since $S$ is a self minimizer of $f$, we have by Lemma 10 that $S$ is contained in some minimizer $S^*$ of $f$. If $S \neq S^*$, then by Lemma 9, $S^* - S$ is a nonempty self minimizer of $f_{V/S}$, a contradiction. Thus, $S = S^*$, as desired. $\square$

**Lemma 15** *Let $\{C, A, D\}$ be a partition of $V$ in which $A \neq \emptyset$; $A$ is prime with respect to $f_{V/C}$; and $f_{V/C}(A) \leq 0$. If $f(A) > 0$, then there is a partition of $A$ into a sequence $A_1, \ldots, A_k$, such that, for each $i \in \{1, \ldots, k\}$, $A_i$ is prime with respect to $f_{V/[A_i)}$, and $f_{V/[A_i)}(A_i) > 0$.*

PROOF. From Theorem 12, it suffices to show that $A$ is a minimizer with respect to $\bar{f}_{V/C \cup D}$. To show this, it suffices to show that $\bar{m}_{V/C \cup D}(B, A - B) = f(B) > 0$, for every $\emptyset \neq B \subset A$.

$$f(B) = m(B, \emptyset) \tag{21}$$
$$\geq m(B, C) \tag{22}$$
$$> 0, \tag{23}$$

where (21) follows by definition; (22) follows by submodularity; and (23) follows because $A$ is prime with respect to $f_{V/C}$ $\square$

8

# 4 Minimal self minimizers

## 4.1 Decomposition of minimal self minimizers

The goal is to efficiently recognize minimal self minimizers. Toward this goal, we have a decomposition result.

**Theorem 16** *Let $A$ be a minimal self minimizer of $f$ with $|A| > 1$. Then there is a partition of $A$ into a sequence of at least two nonempty sets $\{A_1, \cdots, A_k\}$, such that the following hold:*

1. *Each $A_i$ is prime with respect to $f_{V/[A_i)}$, for each $i \in \{1, \ldots, k\}$;*

2. *$f(A_1) > 0$;*

3. *$m(A_k, [A_k)) < 0$;*

4. *$m(A_i, [A_i)) = 0$, for each $i \in \{2, \ldots, k-1\}$;*

5. *if $k > 2$, then $m(A_k, [A_i)) \geq 0$, for each $i \in \{1, \ldots, k-1\}$.*

PROOF. Let $A_1$ be a nonempty proper subset of $A$ with smallest value of $f$, among nonempty proper subsets of $A$, and subject to this, choose $A_1$ to be minimal. Recursively, starting with $i = 2$, as long as $A - [A_i) \neq \emptyset$, define $A_i$ to be a minimal self minimizer of $f_{V/[A_i)}$. Each $A_i$ exists, since $m(A - [A_i), [A_i)) < 0$.

By choice of $A_1$, we have that both $A_1$ and $A - A_1$ are nonempty. Then, since $A$ is a minimal self minimizer, and is therefore prime, we know that $f(A_1) > 0$ and $f_{V/[A_k)}(A_k) < 0$. The fact that $m(A_i, [A_i)) \leq 0$, for each $i \in \{2, \ldots, k-1\}$, follows from the fact that each $A_i$ to be a minimal self minimizer of $f_{V/[A_i)}$. If $m(A_i, [A_i)) < 0$, for some $i$, then this contradicts the choice of $A_1$.

Finally, assume $k > 2$, and let $i \in \{1, \ldots, k-1\}$. If $i = 1$, then $m(A_k, \emptyset) > 0$ follows because $A$ is a minimal self minimizer. Assume $i \in \{2, \ldots, k-1\}$. Supppose $m(A_k, [A_i)) < 0$. Then

$$
\begin{aligned}
f(A_k \cup [A_i)) &= f([A_i)) + m(A_k, [A_i)) \\
&= f(A_1) + m(A_k, [A_i)) \\
&< f(A_1),
\end{aligned}
$$

contradicting choice of $A_1$. □

## 4.2 Composition of minimal self minimizers

We seek a composition theorem for minimal self minimizers. Particularly nice would be a theorem analogous to the decomposition Theorem 16. Such a result would lead directly to a recognition algorithm, since each item in that theorem can be efficiently verified.

We do not have such a result, but we have some observations in that direction.

Throughout this section, let us assume the following. We have a subset $\emptyset \neq A \subset V$, such that $f(A) < 0$, and a partition of $A$ into a sequence $A_1, \ldots, A_p$ of $p \geq 2$ nonempty sets.

The goal is to find conditions under which we can conclude $A$ is a minimal self minimizer. Equivalently, by Lemma 6, we need to show $f(B) > 0$, for each $\emptyset \neq B \subset A$.

Let $\emptyset \neq B \subset A$. Define $B_i := B \cap A_i$, and let $C_i := C \cap A_i$, for each $i \in \{1 \ldots p\}$.

The following theorem is the best we do in this paper, with the conclusion that $A$ is a self minimizer (not necessarily minimal). This theorem also states that $f(B) > 0$ for some subsets $B$ of $A$, and that if $f(B) \leq 0$, then $B$ must contain $A_p$.

**Theorem 17** *Assume $\emptyset \neq A \subset V$, such that $f(A) < 0$, and a partition of $A$ into a sequence $A_1, \ldots, A_p$ of $p \geq 2$ nonempty sets, with the properties:*

    *1. $A_i$ is prime with respect to $f_{V/[A_i)}$, for each $i \in \{1 \ldots p\}$;*

    *2. $f(A_i) > 0$ for each $i \in \{1, \ldots, p\}$;*

    *3. $m(A_1, A_p) < 0$, and $m(A_p, [A_p)) < 0$;*

    *4. $m(A_i, [A_i)) = 0$, for each $i \in \{2, \ldots, p-1\}$;*

*Then $A$ is a self minimizer with respect to $f$.*

*Let $\emptyset \neq B \subset A$. If $A_p - B \neq \emptyset$, then $f(B) > 0$. In addition, there is a set $A' \subseteq A$ such that $A_p \subset A'$, and $A'$ is a minimal self minimizer with respect to $f$. Moreover, every minimal self minimizer with respect to $f$ contained in $A$ properly contains $A_p$.*

PROOF. We leave the fact that $A$ is a self minimizer for last.

Let $\emptyset \neq B \subset A$, with $A_p - B = C_p \neq \emptyset$. Then we have that $m(B_p, [A_{p-1})) \geq 0$. Now,

$$
\begin{align}
f(B) &= \sum_{i \in \{1, \ldots, p\}} m(B_i, B_1 \cup \cdots \cup B_{i-1}) \tag{24} \\
&\geq \sum_{i \in \{1, \ldots, p\}} m(B_i, A_1 \cup \cdots \cup A_{i-1}) \tag{25} \\
&\geq 0, \tag{26}
\end{align}
$$

where (24) follows by definition of marginal value; (25) follows by submodularity; and (26) follows by the facts that $A_i$ is prime with respect to $f_{V/[A_{i-1})}$, $m(A_i, [A_i)) \geq 0$, for $i \in \{1, \ldots, p-1\}$, and $m(B_p, [A_{p-1})) \geq 0$.

Now, to argue $f(B) > 0$, we need to show that at least one term in (24) is strictly positive. If $B_1 \neq \emptyset$, then the first term is positive. Suppose $B_1 = \emptyset$. Let $2 \leq q \leq p$ be the smallest index with $B_q \neq \emptyset$. Then the first nonempty term in (24) is $m(B_q, \emptyset)$. If $B_q \neq A_q$, then this term is positive, since $A_p$ is prime. If $B_q = A_q$, then this term is positive, by item 2.

Next we investigate $A'$. Since $f(A) < 0$, $A$ contains a minimal self minimizer, $A'$. Since $f(B) > 0$, for every nonempty subset that doesn't contain $A_p$, it must be that $A_p \subseteq A'$. Since $f(A_p) > 0$, $A_p \neq A'$.

Finally, we show $A$ is a minimizer by showing $m(A - B, B) \leq 0$, for each $B \subseteq A$. If $A_p - B \neq \emptyset$, this follows easily:

$$
m(A - B, B) = f(A) - f(B) \leq 0,
$$

10

T since $f(A) < 0$ and $f(B) > 0$. Now assume $A_p \subseteq B$. Then:

$$m(A - B, B) = m(C_1, B) + \sum_{i \in \{2,...,p-1\}} m(C_i, B \cup [A_i]) \qquad (27)$$

$$\leq m(C_1, B_1 \cup A_p) + \sum_{i \in \{2,...,p-1\}} m(C_i, B_i \cup [A_i]) \qquad (28)$$

$$\leq 0, \qquad (29)$$

where (27) follows by definition of marginal value, and (28) follows by submodularity. For (29), we show each term of (28) is nonnegative. The first term follows, if $C_1 \neq A_1$, since $A_1$ is prime. If $C_1 = A_1$, then it follows from item 3. That each term in the sum of the second term of (28) is nonnegative follows by item 1, if $C_i \neq A_i$, and by item 4, otherwise. $\square$

We would need at least one more (as yet unknown) condition, besides those in Theorem 17, to prove $A$ is a minimal self minimizer. Indeed, here is an example where $A = A_1 \cup A_2$ is not a minimal self minimizer, although $A$, $A_1$, and $A_2$ satisfy the conditions of Theorem 17. In the example below, $A_1 = \{a, b\}$, $A_2 = \{c\}$.

The needed condition must rule out subsets $B$ with $A_k \subset B \subset A$ and $f(B) < 0$.

| $B$ | $f(B)$ | $m(a,B)$ | $m(b,B)$ | $m(c,B)$ | $m(B,a)$ | $m(B,b)$ | $m(B,c)$ |
|---|---|---|---|---|---|---|---|
| $\{a\}$ | 2 | | -1 | -1 | | -1 | -1 |
| $\{b\}$ | 2 | -1 | | -3 | -1 | | -3 |
| $\{c\}$ | 2 | -1 | -2 | | -1 | -3 | |
| $\{a,b\}$ | 1 | | | -3 | | | -4 |
| $\{b,c\}$ | -1 | -1 | | | -4 | | |
| $\{a,c\}$ | 1 | | -3 | | | -4 | |
| $\{a,b,c\}$ | -2 | | | | | | |

Clearly $A$ is not a minimal self minimizer, since $f(\{b, c\}) < 0$. $A$ is, however, a self minimizer.

In the special case where $p = 2$ and $|A_1| = 1$, we can prove a composition theorem.

**Theorem 18** *Assume a partition $\{E, A, D\}$ of $V$, such that $\emptyset \neq A \subseteq V$. Assume also that $u \in A$, and $A_1 = A - u$, with the property that $A_1$ is prime with respect to $f_{V/(E+u)}$. Assume also the following conditions:*

1. *$f_{V/E}(A) < 0$;*

2. *$f_{V/E}(u) > 0$;*

3. *$f_{V/E}(A_1) > 0$;*

*Then $A$ is prime with respect to $f_{V/E}$.*

PROOF. By item 1 and Lemma 6, we need only to show $f(B) > 0$, for each $\emptyset \neq B \subset A$. To that end, let $\emptyset \neq B \subset A$.

If $B = \{u\}$, then $f(B) > 0$, by item 2. If $B = A_1$, then $f(B) > 0$, by item 3. If $u \in B$, then

$$f(B) = f(u) + m_{V/E}(B \cap A_1, u) \qquad (30)$$

$$> 0, \qquad (31)$$

11

where (30) follows by definition of marginal value; and (31) follows by item 2 and the fact that $A_1$ is prime with respect to $f_{V/(E+u)}$.

Otherwise, $B \subset A$, and:

$$f(B) \;=\; m_{V/E}(B \cap A_1, \emptyset) \tag{32}$$
$$\;\geq\; m_{V/E}(B \cap A_1, u) \tag{33}$$
$$\;>\; 0, \tag{34}$$

where (32) follows by definition of marginal value; (33) follows by submodularity. (34) follows by item 2 and the fact that $A_1$ is prime with respect to $f_{V/(E+u)}$. $\square$

Lest one might believe a similar theorem holds in the case that item 1 is replaced with $f_{V/E}(A) > 0$, we have this counterexample shown in the table below, where $V = \{a, b, c\}$.

It is straightforward to check that letting $A = \{a, b, c\}$ and $A_1 = \{b, c\}$, we have that $A_1$ is prime with respect to $f_{V/a}$; $f(a) > 0$; $f(A_1) > 0$; and $f(A) > 0$. However, $A$ is not prime with respect to $f$. Indeed, $m(\{a, b\}, c) > 0$.

| $B$ | $f(B)$ | $m(a,B)$ | $m(b,B)$ | $m(c,B)$ | $m(B,a)$ | $m(B,b)$ | $m(B,c)$ |
|---|---|---|---|---|---|---|---|
| $\{a\}$ | 1 | | 2 | 2 | | 0 | 0 |
| $\{b\}$ | 1 | 2 | | 2 | 0 | | 0 |
| $\{c\}$ | 1 | 2 | 2 | | 0 | 0 | |
| $\{a, b\}$ | 3 | | | -1 | | | 2 |
| $\{b, c\}$ | 3 | -1 | | | 2 | | |
| $\{a, c\}$ | 3 | | -1 | | | 2 | |
| $\{a, b, c\}$ | 2 | | | | | | |

$\square$

Another possible analog would be the following conditions:

1. $A_1$ prime with respect to $f_{V/(E+u)}$;

2. $f_{V/E}(A) > 0$;

3. $m_{V/E}(A_1, u) < 0$;

4. $m_{V/E}(u, A_1) < 0$;

But again, these don't imply $A$ is prime with respect $f_{V/E}$, as the following example shows:

| $B$ | $f(B)$ | $m(a,B)$ | $m(b,B)$ | $m(c,B)$ | $m(B,a)$ | $m(B,b)$ | $m(B,c)$ |
|---|---|---|---|---|---|---|---|
| $\{a\}$ | 3 | | 1 | 1 | | 1 | 1 |
| $\{b\}$ | 3 | 1 | | -2 | 1 | | -2 |
| $\{c\}$ | 3 | 1 | -2 | | 1 | -2 | |
| $\{a, b\}$ | 4 | | | -2 | | | -1 |
| $\{b, c\}$ | 1 | 1 | | | -1 | | |
| $\{a, c\}$ | 4 | | -2 | | | -1 | |
| $\{a, b, c\}$ | 2 | | | | | | |

In this example, $A = V = \{a, b, c\}$. Let $u = a$, and $A_1 = \{b, c\}$. It is straightforward to check that $A_1$ is prime with respect to $f_{V/a}$ and the conjectured conditions hold. But since $m(a, \{b, c\}) > 0$, the set $A$ cannot be prime. $\square$

Another possible conjecture is that every prime set decomposes as in Theorem 18. Unfortunately this is not true. Below is an example of a minimal self minimizer $A$ such that $A - u$ is not a minimal self minimizer with respect to $f_{V/u}$, for every $u \in V$.

| $B$ | $f(B)$ | $m(a,B)$ | $m(b,B)$ | $m(c,B)$ | $m(B,a)$ | $m(B,b)$ | $m(B,c)$ |
|---|---|---|---|---|---|---|---|
| $\{a\}$ | 2 | | -1 | -1 | | -1 | -1 |
| $\{b\}$ | 2 | -1 | | -1 | -1 | | -1 |
| $\{c\}$ | 2 | -1 | -1 | | -1 | -1 | |
| $\{a,b\}$ | 1 | | | -2 | | | -3 |
| $\{b,c\}$ | 1 | -2 | | | -3 | | |
| $\{a,c\}$ | 1 | | -2 | | | -3 | |
| $\{a,b,c\}$ | -1 | | | | | | |

□

# 5    Pseudoprime Sets

In Sections 4.1 and 4.2, it becomes evident that sets with zero marginal value add complication. For this reason, we introduce the notion of pseudoprime sets. A psuedoprime set $A$ is a prime set unless it has a nonempty proper subset with function value equal to 0 or to that of set $A$ itself.

In this section, we define pseudoprime sets and list the properties and results that are analogous to those for prime sets. In each case, the proof is similar but simpler to its prime-set analog. Therefore, we omit these proofs.

Given a normalized submodular function $f$ on $V$, a set $A \subseteq V$ is *pseudoprime* if for each $\emptyset \neq B \subset A$, the following two properties hold:

- $f(B) \geq 0$, and

- $m(B, A - B) \leq 0$.

Pseudoprime sets have many of the same properties as prime sets, and we list several here without proof. For this purpose, assume $V$ is partitioned into a sequence $\{C, A, D\}$.

- $A$ is pseudoprime with respect to $f_{V/C}$ if and only if $A$ is pseudoprime with respect to $\bar{f}_{V/D}$.

- If $f_{V/C}(A) \leq 0$, then $A$ is pseudoprime with respect to $f_{V/C}$ if and only if $f_{V/C}(B) \geq 0$, for every $\emptyset \neq B \subset A$.

- If $f_{V/C}(A) \geq 0$, then $A$ is pseudoprime with respect to $f_{V/C}$ if and only if $m_{V/C}(B, A - B) \leq 0$, for every $\emptyset \neq B \subset A$.

- Assume A is pseudoprime with respect to $f_{V/C}$, and $f_{V/C}(A) \leq 0$. If $f(A) \leq 0$, then $A$ is pseudoprime with respect to $f$.

The connection between pseudoprime sets and minimal self minimizers is not very natural. However, we do have the following decomposition of minimizers into pseudoprime self minimizers.

- If $A$ is pseudoprime with respect to $f_{V/C}$ and $f_{V/C}(A) \leq 0$, then $A$ is a self minimizer with respect to $f_{V/C}$.

- If $A$ is pseudoprime and $f_{V/C}(A) \geq 0$, then $A$ is a self minimizer with respect to $\bar{f}_{V/C}$.

- If $A$ is a self minimizer with respect to $f$, then $A$ can be partitioned into a sequence $\{A_1, A_2, \ldots, A_k\}$, such that for each $i \in \{1, \ldots, k\}$, $A_i$ is pseudoprime with respect to $f_{V/[A_i)}$ and $f_{V/[A_i)}(A_i) \leq 0$.

- If $A$ is partitioned into a sequence $\{A_1, A_2, \ldots, A_k\}$, such that for each $i \in \{1, \ldots, k\}$, $A_i$ is pseudoprime with respect to $f_{V/[A_i)}$ and $f_{V/[A_i)}(A_i) \leq 0$, then $A$ is a self minimizer with respect to $f$.

**Theorem 19** *There is a partition of $V$ into a sequence of sets*

$$\{S_1, S_2, \ldots, S_k, T_1, T_2, \ldots, T_l\},$$

*such that*

- $f_{V/[S_i)}(S_i) \leq 0$, and $S_i$ is pseudoprime with respect to $f_{V/[S_i)}$, for each $i \in \{1, \ldots, k\}$;

- $f_{V/[T_i)}(T_i) \geq 0$, and $T_i$ is pseudoprime with respect to $f_{V/[T_i)}$, for each $i \in \{1, \ldots, l\}$.

*Given such a partition, the set $S := S_1 \cup S_2 \cup \cdots \cup S_k$ is a minimizer of $f$. Moreover, every minimizer of $f$ decomposes in this way.*

□

A sequence of the type ensured by Theorem 19 is called a *pseudoprime decomposition* of $V$.

The decomposition theorem for pseudoprime sets that is analagous to Theorem 16 is simpler:

**Theorem 20** *Let $A$ be a pseudoprime set with respect to $f$. Then there is a partition of $A$ into a sequence of two nonempty sets $\{A_1, A_2\}$, such that the following hold:*

- $A_1$ *is pseudoprime with respect to $f$ and $A_2$ is pseudoprime with respect to $f_{V/A_1}$; $f(A_1) \geq 0$ and $m(A_2, A_1) \leq 0$;*

- $f(A_2) \geq 0$ *and $m(A_1, A_2) \leq 0$;*

As for composition, as in the case of primes, the conditions under which a partition of $A$ into sequence $\{A_1, A_2\}$ satisfying the above three properties implies $A$ is pseudoprime remain an open question. We do, however, have an analog to Theorem 17.

**Theorem 21** *Assume $\emptyset \neq A \subset V$, such that $f(A) \leq 0$, and a partition of $A$ into a sequence of nonempty sets $\{A_1, A_2\}$, with the properties:*

1. $A_1$ *is pseudoprime with respect to $f$ and $A_2$ is pseudoprime with respect to $f_{V/A_1}$;*

2. $f(A_1) \geq 0$ *and $m(A_2, A_1) \leq 0$;*

*3. $f(A_2) \geq 0$ and $m(A_1, A_2) \leq 0$;*

*Then $A$ is a self minimizer with respect to $f$.*

*We do give a polynomial certificate for verifying a set $A$ is a pseudo-prime set in Theorem 25. Unfortunately, this is a convex certificate involving not just one but $|A|$ vectors.*

*Let $\emptyset \neq B \subset A$. If $A_2 - B \neq \emptyset$, then $f(B) \geq 0$. In addition, if $f(A) < 0$, then there is a set $A' \subseteq A$ such that $A_2 \subset A'$, $f(A') < 0$, and $A'$ is pseudoprime with respect to $f$. Moreover, every such subset $A'$ of $A$ properly contains $A_2$.*

# 6  Polyhedral and Algorithmic Issues

Given a submodular funcion $f$ on $V$, with $f(\emptyset) = 0$, Edmonds [4] defined the polyhedron:

$$P_f := \{x \in \mathbf{R}^V \ : \ x(V) = f(V); \ x(A) \leq f(A), \text{ for every } A \subseteq V\},$$

where $\mathbf{R}^V$ denotes the $|V|$-dimensional real vector space with components indexed on $V$, and for $x \in \mathbf{R}^V$ and $A \subseteq V$, $x(A) := \sum_{v \in A} x_v$. Given a total order $\prec$ of the set $V$, the greedy vector $b^\prec$ is defined as follows:

$$b_v^\prec := f([v]^\prec) - f([v)^\prec).$$

Here are 2 definitions equivalent to the one above:

$$b_v^\prec \ := \ m(v, [v)^\prec). \tag{35}$$
$$b_v^\prec \ := \ f_{V/[v)^\prec}(v). \tag{36}$$

**Theorem 22 (Edmonds [4])** *The extreme points of $P_f$ are exactly the greedy vectors $b^\prec$.*

$\square$

Given a partition of $V$ into a sequence $\{E, A, D\}$ of subsets, define $f_{V/E \backslash D}$ to be the submodular function on $A$, where for $B \subseteq A$,

$$f_{V/E \backslash D}(B) := f(B \cup E) - f(E).$$

Given vectors $x \in \mathbf{R}^E$ and $y \in \mathbf{R}^A$, let $x|y \in \mathbf{R}^{E \cup A}$ be the concatination of $x$ and $y$.

**Proposition 23** *Let $\{E, A\}$ be a partition of $V$, $x \in P_{V \backslash A}$, and $y \in P_{f_{V/E}}$. Then $x|y \in P_f$.*

PROOF. For convenience, let $z = x|y$. First we show $z(V) = f(V)$. We have

$$\begin{aligned} z(V) \ &= \ x(E) + y(A) \\ &= \ f(E) + f_{V/E}(A) \\ &= \ f(E) + f(E \cup A) - f(E) \\ &= \ f(V), \end{aligned} \tag{37}$$

as desired. Now let $\emptyset \neq B \subset V$.

$$
\begin{aligned}
z(B) &= x(B \cap E) + y(B \cap A) \\
&\leq f(B \cap E) + f_{V/E}(B \cap A) \\
&= f(B \cap E) + m(B \cap A, E) \\
&\leq f(B \cap E) + m(B \cap A, B \cap E) \qquad (38) \\
&= f(B), \qquad\qquad\qquad\qquad\qquad\qquad (39)
\end{aligned}
$$

where (38) follows from submodularity of $f$, and (39) follows from the definition of marginal value. $\square$

**Corollary 24** *The greedy vector $b^{\prec} \in P_f$.*

$\square$

The vector that is 1 in component $v$ and zero elsewhere is denoted $\chi_v$. The dimension of $\chi_v$ should always be clear from the context.

**Theorem 25** *Assume a normalized submodular function $f$ on $V$, and $\emptyset \neq A \subseteq V$. Then $A$ is pseudoprime with respect to $f$ if and only if the vector $f(A)\chi_v \in P_{f_{V \setminus (V-A)}}$, for every $v \in A$.*

PROOF. ($\Longrightarrow$) Assume $\emptyset \neq A$ is pseudoprime with respect to $f$. Let $v \in A$, and let $\emptyset \neq B \subseteq A$. If $v \in B$, then
$$
f(A)\chi_v(B) = f(A) \leq f(B),
$$
since $A$ is a self minimizer, as desired. If $B = A$, then equality holds, as required. Assume $v \notin B$. Then
$$
f(A)\chi_v(B) = 0 \leq f(B),
$$
since $A$ is pseudoprime, as desired. $\square$

($\Longleftarrow$) Assume $f(A)\chi_v \in P_f$, for every $v \in A$. To show $A$ is pseudoprime, let $\emptyset \neq B \subset A$. Let $v \in A - B$. Since $f(A)\chi_v \in P_f$, we have
$$
0 = f(A)\chi_v(B) \leq f(B),
$$
as desired. We also have
$$
f(A) = f(A)\chi_v(A - B) \leq f(A - B),
$$
which gives us
$$
m(B, A - B) = f(A) - f(A - B) \leq 0,
$$
as desired. $\square$

**Theorem 26** *Let $\{S_1, S_2, \ldots, S_p, T_1, T_2, \ldots, T_q\}$ be a pseudoprime decomposition of $V$. Then there exists a collection $W = \{(\lambda_i, \prec_i, b^i) : i = 1, \ldots, k\}$ of weighted extreme vectors, where for each $i = 1, \ldots, k$, $b^i$ is the greedy vector corresponding to $\prec_i$ and $\lambda_i > 0$; $\sum_{i=1}^k \lambda_i = 1$; and vector $x := \sum_{i=1}^k \lambda_i b^i$, such that:*

16

*1. $x_v \leq 0$, for each $v \in S_1 \cup \cdots \cup S_p$;*

*2. $x_v \geq 0$, for each $v \in T_1 \cup \cdots \cup T_q$);*

*3. for each $j \in \{1, \ldots, p\}$, the set $(S_j]$ is a lower ideal of each order $\prec_i \in W$;*

*4. $|\{v \in S_j \ : \ x_v \neq 0\}| \leq 1$, for each $j \in \{1, \ldots, p\}$;*

*5. for each $j \in \{1, \ldots, p\}$, the set $(T_j]$ is a lower ideal of each order $\prec_i \in W$;*

*6. $|\{v \in T_j \ : \ x_v \neq 0\}| \geq 1$, for each $j \in \{1, \ldots, q\}$.*

PROOF. This is straightforward from Theorem 25 and Proposition 23. □

Next we show how to find a pseudoprime decomposition, together with the collection $W$ of weighted extreme vectors guaranteed by Theorem 26. Our approach involves modifying Schrijver's algorithm for submodular function minimization, which we review next.

## 6.1 Schrijver's Algorithm

Here we review Schrijver's polynomial algorithm for minimizing a submodular function. Given a collection $\{\prec_1, \prec_2, \ldots, \prec_k\}$ of total orders on $V$, the digraph $G = (V, A)$ has an arc $(u, v) \in A$ if $u \prec_i v$, for some $i \in \{1, \ldots, k\}$. Given total order $\prec$ and elements $s \prec t \in V$, the interval $(s, t]_\prec := \{u \in V \ : \ s \prec u \preceq t\}$.

**Input:** A submodular function $f : 2^V \rightarrow \Re$, having $f(\emptyset) = 0$, given via an oracle.

**Output:**   1. A collection of weighted extreme points:
$W = \{(\lambda_1, \prec_1, b^{\prec_1}), (\lambda_2, \prec_2, b^{\prec_2}), \ldots, (\lambda_k, \prec_k, b^{\prec_k})\}$, where $\sum_{i=1}^k \lambda_i = 1$, and for each $i \in \{1, \ldots, k\}$, $\lambda_i \geq 0$, $\prec_i$ is a total order of $V$, and $b^{\prec_i}$ is the vector obtained via the greedy algorithm and order $\prec_i$.

  2. A set $U^* \subseteq V$ that is a lower ideal of each order $\prec_1, \prec_2, \ldots, \prec_k$, and such that, where $x := \sum_{i=1}^k \lambda_i b^{\prec_i}$, the set $U^*$ contains every element $v \in V$ with $x_v < 0$ and $U^*$ contains no element $v \in V$ with $x_v > 0$.

**Step 0:** [Initialize]
Let $\prec_1$ be an arbitrary order of $V$, let $b^{\prec_1}$ be the corresponding extreme point, and let $\lambda_1 = 1$. Put $W := \{(\lambda_1, \prec_1, b^{\prec_1})\}$. Index the elements of $V = \{1, 2, \ldots, n\}$.

**Step 1:** [Construct Auxiliary Digraph with Distances]
Create auxiliary directed graph $G = (V, A)$, where
$A := \{(u, v) : u \prec_i v, \text{ for some } i = 1, \ldots, n\}$. Define $x := \sum_{i=1}^k \lambda_i b^{\prec_i}$.
Define $P := \{v \in V : x_v > 0\}$, and $N := \{v \in V : x_v < 0\}$.
For each $v \in V$, define $d_v$ to be the number of arcs in a shortest dipath to $v$ from some node in $P$.

**Step 2:** [Check Optimality Condition]
If $d_v = +\infty$, for each $v \in N$, then let
$U^* := \{v \in V : \text{there is a dipath from } v \text{ to some node in } N\}$.
Return $W$ and $U^*$.

**Step 3:** [Choose $s$, $t$, and $\prec_1$]

Choose $t \in N$ with $d_t$ maximum among $v \in V$ with finite $d_v$, breaking any ties by choosing $t$ with maximum index. Choose $s \in V$ such that $(s,t) \in A$, $d_s = d_t - 1$, and $s$ having maximum index. Choose order $\prec_1$ (renumbering if necessary) with $s \prec_1 t$ and with the largest $|(s,t]_{\prec_i}|$.

**Step 4:** [Improve]

Compute $\delta \geq 0$ such that $b^{\prec_1} + \delta(\chi^t - \chi^s)$ is a convex combination of the extreme points $b^{\prec_1^{s,u}}$, for $u \in (s,t]_{\prec_1}$.

Compute $y = x + \lambda_1 \delta(\chi^t - \chi^s)$.

Let $x'$ be the point on the line segment $\overline{xy}$ closest to $y$ and having $x'_t \leq 0$. Describe $x'$ as a convex combination of $b^{\prec_i}$, for $i = 1, \ldots, k$, and $b^{\prec_1^{s,u}}$, for $u \in (s,t]_{\prec_1}$, where either $x'_t = 0$ or $b^{\prec_1}$ is removed from the list. Update $W$.

**Step 5:** [Reduce W]

Apply Gaussian elimination to reduce the set $W$ of weighted extreme points to one having size at most $|V|$. Return to Step 1.

## 6.2 Correctness and Complexity of Schrijver's Algorithm

Schrijver shows the number of iterations is bounded by $|V|^6$ as follows. In the current iteration, let

$$\alpha := |(s,t]_{\prec_1}|,$$

and let

$$\beta := \text{ number of } i \in \{1, \ldots, k\} \text{ with } |(s,t]_{\prec_i}| = \alpha.$$

Let $x'$, $d'$, $A'$, $P'$, $N'$, $t'$, $s'$, $\alpha'$, $\beta'$ be the objects $x$, $d$, $A$, $P$, $N$, $t$, $s$, $\alpha$, $\beta$ in the next iteration.

For each iteration of the algorithm $d'_v \geq d_v$, for each $v \in V$. (This relies on the fact that $P' \subseteq P$ and that the only arcs in $A' \setminus A$ are of the type $(u,v)$, where $s \preceq_1 v \prec_1 u \preceq_1 t$.) Moreover, if after some iteration, $d'_v = d_v$, for each $v \in V$, then the vector $(d'_{t'}, t', s', \alpha', \beta')$ is lexicographically smaller than $(d_t, t, s, \alpha, \beta)$. While $d_v$ remains constant, for each $v \in V$, there are at most $|V|$ different possible values for $s$, $\alpha$, and $\beta$, and at most $|V|$ different possible pairs $(d_t, t)$. Thus, there are at most $|V|^4$ iterations between two consecutive iterations in which some $d_v$ increases. Finally, there can be at most $|V|^2$ iterations in which some $d_v$ increases.

Given that the algorithm terminates, it must do so in Step 2. At that point, since $U^*$ has no incoming arcs in $G$, it follows that $U^*$ is a lower ideal in each $\prec_i$, $i \in \{1, \ldots, k\}$. Moreover, $U^*$ contains every member of $N$ and no member of $P$. It now follows that for every $S \subseteq V$,

$$
\begin{align}
f(U^*) &= x(U^*) \tag{40}\\
&\leq x(S) \tag{41}\\
&\leq f(S), \tag{42}
\end{align}
$$

where (40) follows from the fact that $U^*$ is a lower ideal of $\prec_i$, for each $\imath \in \{1, \ldots, k\}$; (41) follows since $U^*$ contains every negative component and no positive component of $x$; and (42)

follows because $x \in P_f$. This proves optimality of $U^*$ and completes the proof of correctness.
$\square$

## 6.3 Strong Components of $G$

In this section, we show that Schrijver's algorithm can be modified to construct a pseudoprime decomposition of $V$, together with the collection $W$ and vector $x$ guaranteed by Theorem 26. Our analysis shows the complexity of this modified algorithm is $O(|U^*|^8)$.

First we have some definitions about digraphs.

Given a digraph $G = (V, A)$, and $U \subseteq V$, let $G[U] := (U, A')$ be the subgraph of $G$ induced by $U$, where arc $(u, v) \in A'$ if and only if $u \in U$, $v \in U$, and $(u, v) \in A$.

Digraph $G$ is *strongly connected* if for every pair of nodes $u, v \in V$, $G$ has a dipath from $u$ to $v$.

For every digraph $G$ there is a unique partition of $V$ into a sequence $\{V_1, V_2, \ldots, V_r\}$, where for each $i \in \{1, \ldots, r\}$,

- $G_i := G[V_i]$ is strongly connected; and

- any arc $(u, v)$ with $u \in V_i$ has $v \in V_j$, for some $j \geq i$.

The digraphs $G_i$ are called the *strong components* of $G$.

During execution, Schrijver's algorithm constructs and maintains an auxiliary digraph. Let $G = (V, A)$ be this digraph at the time the algorithm terminates. Let us also have the final subsets $U^*$, $P$, and $N$ of $V$, the final vector $x$, together with the collection of weighted extreme points:

$$W = \{(\lambda_1, \prec_1, b^{\prec_1}), (\lambda_2, \prec_2, b^{\prec_2}), \ldots, (\lambda_k, \prec_k, b^{\prec_k})\}$$

where $\sum_{i=1}^{k} \lambda_i = 1$, and for each $i \in \{1, \ldots, k\}$, $\lambda_i \geq 0$, $\prec_i$ is a total order of $V$, and $b^{\prec_i}$ is the vector obtained via the greedy algorithm and order $\prec_i$.

We add a post-processing phase to Schrijver's algorithm, the purpose of which is to reach the stage where each strong component of $G[U^*]$ has at most one node $v$ for which $x_v \neq 0$.

## Strong Components Algorithm

**Input:** A submodular function $f : 2^V \rightarrow \Re$, having $f(\emptyset) = 0$, given via an oracle; the output from Schrijver's Algorithm applied to $f$.

**Output:**   1. A collection of weighted extreme points:
$W = \{(\lambda_1, \prec_1, b^{\prec_1}), (\lambda_2, \prec_2, b^{\prec_2}), \ldots, (\lambda_k, \prec_k, b^{\prec_k})\}$, where $\sum_{i=1}^k \lambda_i = 1$, and for each $i \in \{1, \ldots, k\}$, $\lambda_i \geq 0$, $\prec_i$ is a total order of $V$, and $b^{\prec_i}$ is the vector obtained via the greedy algorithm and order $\prec_i$.

2. Sets $S_1, S_2, \ldots, S_p$, such that:
  (a) each set $S_i$ is a strong component of $G$;
  (b) each set $S_1 \cup \ldots \cup S_i$ is a lower ideal of each $\prec_j$;
  (c) for each set $S_i$, the vector $x$ has at most one nonzero component and that component is $x_v \leq 0$.

**Step 0:** [Initialize]

Delete $V - U^*$ from $G$ and from each member of $W$.

Let $S_1, S_2, \ldots, S_p$ be a sequenced partition of $U^*$ into the node sets of the strong components of $G$, ordered so that each set $(S_i]$ is a lower ideal of each order in $W$. Index the members of $U^*$ consistent with the sequence $S_1, S_2, \ldots, S_p$.

For each $i \in \{1 \ldots p\}$, choose one root node $u_i \in S_i$.

For each $i \in \{1 \ldots p\}$ let $G_i = (S_i, A_i)$ be $G$ restricted to node set $S_i$.

**Step 1:** [Compute Distances]

For each $i \in \{1 \ldots p\}$, define $d_u$ to be the number of arcs in a shortest dipath to $u$ from node $u_i$. If some node $u \in S_i$ is not reachable from $u_i$, then subdivide $S_i$ into its sequenced partition of strong components, and choose a root for each component that does not contain the current $u_i$. Renumber the sets $S_i$ to reflect this subdivision.

**Step 2:** [Check Doneness]

If $x_u = 0$, for each $u \in S_i - u_i$, and each $i \in \{1 \ldots p\}$, and if every element $v \in V$ has at some point been a designated $u_i$, then stop and return $S_1, S_2, \ldots, S_p$. If $x_u = 0$, for each $u \in S_i - u_i$, and each $i \in \{1 \ldots p\}$, but some $v \in V$ has not been a designated $u_i$ for the strong component to which it currently belongs, then choose such a $v \in S_i$, say, and designate $v$ as $u_i$.

**Step 3:** [Choose $s$, $t$, in $S_i$ and $\prec_1$]

Choose a component $G_i = (S_i, A_i)$ with some nonzero component other than $x_{u_i}$. Choose $t \in S_i - u_i$ with $x_t < 0$ and $d_t$ maximum among $v \in S_i - u_i$ with $x_t < 0$, breaking any ties by choosing $t$ with maximum index. Choose $s \in S_i$ such that $(s, t) \in A_i$, $d_s = d_t - 1$, and $s$ having maximum index. Choose order $\prec_1$ (renumbering if necessary) with $s \prec_1 t$ and with the largest $|(s, t]_{\prec_i}|$.

**Step 4:** [Improve with $s$ and $t$ in $S_i$]

Compute $\delta \geq 0$ such that $b^{\prec_1} + \delta(\chi^t - \chi^s)$ is a convex combination of the extreme points $b^{\prec_1^{s,u}}$, for $u \in (s, t]_{\prec_1}$.

Compute $y = x + \lambda_1 \delta(\chi^t - \chi^s)$.

Let $x'$ be the point on the line segment $\overline{xy}$ closest to $y$ and having $x'_t \leq 0$. Describe $x'$ as a convex combination of $b^{\prec_i}$, for $i = 1, \ldots, k$, and $b^{\prec_1^{s,u}}$, for $u \in (s,t]_{\prec_1}$, where either $x'_t = 0$ or $b^{\prec_1}$ is removed from the list. Update $W$.

**Step 5:** [Reduce W]

Apply Gaussian elimination to reduce the set $W$ of weighted extreme points to one having size at most $|U^*|$. Return to Step 1.

## 6.4 Correctness and Complexity of Strong Components Algorithm

We show the number of iterations dealing with $G_i$ is bounded by $|U^*|^7$ as follows. In the current iteration, let

$$\alpha := |(s,t]_{\prec_1}|,$$

and let

$$\beta := \text{ number of } i \in \{1, \ldots, k\} \text{ with } |(s,t]_{\prec_i} = \alpha.$$

Let $x'$, $d'$, $A'$, $P'$, $N'$, $t'$, $s'$, $\alpha'$, $\beta'$ be the objects $x$, $d$, $A$, $P$, $N$, $t$, $s$, $\alpha$, $\beta$ in the next iteration.

A subdivision may occur at most $|U^*|$ times. Between 2 iterations involving a subdivision, an iteration involving a reassignment of some $u_i$ may occur at most $|U^*|$ times. Now let us develop an upper bound on the number of iterations between 2 consecutive iterations starting with a subdivision or a reassignment.

For each iteration that does not start with a subdivision or reassignment, $|U^*| \geq d'_v \geq d_v$, for each $v \in U^*$. Thus, there can be at most $|U^*|^2$ iterations in which some $d_v$ increases, between 2 consecutive iterations having subdivisions or reassignment.

Now let us develop an upper bound on the number of consecutive iterations in which $d$ remains constant.

First we argue that $(d'_{t'}, t', s', \alpha', \beta')$ is lexicographically smaller than $(d_t, t, s, \alpha, \beta)$. Assume $d'_v = d_v$ for every $v \in U^*$. If $d'_{t'} < d_t$, then we are done. If not, then by choice of $t$, $d'_{t'} = d_t$. If $t \neq t'$, then $t < t'$, by choice of $t$. Assume $t = t'$. Then $x'_t < 0$, which means $b^{\prec_1}$ was removed from $W$. If $s' \neq s$, then $s' < s$, and we are done. Otherwise, we have $s' = s$ and $t' = t$. Any orders introduced in the iteration have $|(s,t]| < \alpha$. Therefore, since $\prec_1$ was chosen with greatest $|(s,t]|$, it must be that $\alpha' \leq \alpha$. If $\alpha' = \alpha$, then since $\prec_1$ has been removed, we can conclude that $\beta' < \beta$. This completes the lexicographic argument.

There are at most $|U^*|$ different possible values for $\beta$, $\alpha$, and $s$. While $d$ remains constant, there are at most $|U^*|$ different pairs $(d_t, t)$. Thus, there are at most $|S_i|^4$ iterations between two consecutive iterations in which some $d_v$ increases.

Together with the fact that there can be at most $|U^*|^2$ iterations in which some $d_v$ increases, between 2 consecutive iterations with subdivision or reassignment, and the fact that there are an most $|U^*|^2$ iterations with subdivision or reassignment, gives the bound of at most $2|U^*|^8$ iterations.

Given that the algorithm terminates, it must do so in Step 2. At that point, the sequence $S_1, \ldots, S_p$ has the required properties. $\square$

The strong components algorithm works only with the minimizer $U^*$, and returns the decomposition $\{S_1, \ldots, S_k\}$ of $U^*$. It may be applied to $V - U^*$ and $\bar{f}$ to obtain $\{T_1, \ldots, T_l\}$.

## 6.5  Extreme point solution

Consider the following primal-dual pair of linear programming problems:

$$\textbf{P:} \qquad \begin{aligned} \text{Maximize} \quad & x(V) \\ \text{Subject to:} \quad x(A) \;&\le\; f(A), \quad \text{for each } A \subseteq V \\ x_v \;&\le\; 0, \qquad \text{for each } v \in V \end{aligned}$$

$$\textbf{D:} \qquad \begin{aligned} \text{Minimize} \quad & \sum_{A \subseteq V} f(A)\, y_A \\ \text{Subject to:} \quad \sum_{A \subseteq V\,:\,v \in A} y_A \;&\le\; 1, \quad \text{for each } v \in V \\ y_A \;&\ge\; 0, \quad \text{for each } A \subseteq V \end{aligned}$$

Let $x$ and $U^*$ be the vector and minimizing set obtained by executing Schrijver's algorithm. Let $\bar{x}$ be the vector obtained from $x$ by replacing each component $x_v$ with $min(x_v, 0)$. Let $\bar{y} \in \mathbf{R}^{2^V}$ be the vector that is 1 on $U^*$ and 0 elsewhere.

It is known and not difficult to show (by showing that they are feasible and complementary) that $\bar{x}$ and $\bar{y}$ are optimal solutions to **P** and **D**, respectively. Indeed, this is the proof of optimality of $U^*$.

We know for linear programming theory that every linear programming problem that has an optimal solution has an extreme-point optimal solution. One might ask for an extreme-point optimal solution to **P**. We have the following.

**Theorem 27** *Let $x$ be the vector obtained by executing Schrijver's algorithm followed by the Strong Components algorithm. Let $\bar{x}$ be the vector obtained from $x$ by replacing each component $x_v$ with $min(x_v, 0)$. Then $\bar{x}$ is an extreme-point optimal solution to* **P**.

Proof. We already know that $\bar{x}$ is an optimal solution to **P**. To show that $\bar{x}$ is an extreme point, it suffices to provide an ordering $\prec$ of $V$ so that for each $v \in V$, we have either that $x_v = 0$ or $\bar{x}([v)^\prec) = f([v)^\prec)$. Let $\{S_1, S_2, \ldots, S_k\}$ be the pseudo-prime decomposition obtained after execution of Schrijver's algorithm and the Strong Components algorithm, and let $u_i$ be the last designated root for $S_i$, for each $i \in \{1, 2, \ldots, k\}$. Take ordering $\prec$ to start with the elements $S_1 \cup \cdots \cup S_k$, be compatible with the sequence $\{S_1, S_2, \ldots, S_k\}$, and have $u_i$ last within $S_i$, for each $i \in \{1, 2, \ldots, k\}$. The elements of $V - (S_1 \cup \cdots \cup S_k)$ come later, in any order. It is easy to check, using the fact that $\bar{x}_{u_i} = f_{V/[S_i)}(S_i)$, for each $i \in \{1, 2, \ldots, k\}$, that this ordering $\prec$ satisfies the required conditions. $\square$

# 7  Conclusions

We conclude with some questions.

Is there a way to strengthen Theorem 21 or Theorem 17 to conclude that $A$ is prime pseudoprime?

Investigate further the structure of strong components of final digraph $G$ of the Strong Components algorithm. These correspond to the sets in a pseudo-prime decomposition. Let's take the first one, $G_1 = (S_1, A_1)$. For each choice of $u_1 \in S_1$ the algorithm gives a $G_1$. There are possibly many different graphs $G_1$ the algorithm could give for each $u_1$, and also possibly many different still for different choices of $u_1$. Can we say anything about these digraphs? For example, is there one digraph that "works" for all choices of $u_1$?

Can the complexity bound obtained here for the Strong Components algorithm be improved?

# References

[1] R. E. Bixby, W. H. Cunningham, and D. M. Topkis. The Partial Order of a Polymatroid Extreme Point. *Math. of OR* **10** (1985) 367-378.

[2] W. H. Cunningham. Testing Membership in Matroid Polyhedra. *JCT Series B* **36** (1984) 161-188.

[3] W. H. Cunningham . On Submodular Function Minimization. *Combinatorica* **3** (1985) 185-192.

[4] J. Edmonds. Submodular Functions, Matroids, and Certain Polyhedra. *Combinatorial Structures and their Applications*, R. Guy, H. Hanani, N. Sauer, and J. Schonheim, eds., Gordon and Breach, (1970) 6987.

[5] L. K. Fleischer. Recent Progress in Submodular Function Minimization. *Optima*, September 2000 (2000) 1-11.

[6] L. K. Fleischer and S. Iwata. A Push-Relabel Framework for Submodular Function Minimization and Applications to Parametric Optimization. To appear in *Discrete Applied Mathematics* (2001).

[7] L. K. Fleischer and S. Iwata. Improved Algorithms for Submodular Function Minimization and Submodular Flow. *Proceedings of the 32th Annual ACM Symposium on Theory of Computing* (2000) 107-116.

[8] S. Fujishige. *Submodular Functions and Optimization*. North-Holland (1991).

[9] S. Fujishige and S. Iwata. A Descent Method for Submodular Function Minimization. *Math. Prog.* **92** (2002) 387-390

[10] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica* **1** (1981) 169–197. [corrigendum: *Combinatorica* **4** (1984) 291–295].

[11] S. Iwata. A Capacity Scaling Algorithm for Convex Cost Submodular Flows. *Math. Programming* **76** (1997) 299-308.

[12] S. Iwata. A Fully Combinatorial Algorithm for Submodular Function Minimization. *J. Combin. Theory Ser. B* **84** (2002) 203-212.

[13] S. Iwata. A Faster Scaling Algorithm for Minimizing Submodular Functions. *Proceedings of the 9th Conference on Integer Programming and Combinatorial Optimization (IPCO MIT), Lecture Notes in Computer Science 2337*, W. J. Cook and A. S. Schulz eds., Springer, Berlin (2002) 1-8.

[14] S. Iwata, L. Fleischer, and S. Fujishige. A Combinatorial, Strongly Polynomial-Time Algorithm for Minimizing Submodular Functions. *J. ACM* **48** (2001) 761-777.

[15] L. Lovász. Submodular Functions and Convexity. In *Mathematical Programming  The State of the Art*, A. Bachem, M. Grötschel, B. Korte eds., Springer, Berlin (1983) 235-257.

[16] S. T. McCormick. Submodular Function Minimization. A Chapter of the Handbook on Discrete Optimization. Manuscript (2003).

[17] M. N. Queyranne. Minimizing Symmetric Submodular Functions. *Math. Prog.* **82** (1998) 3-12.

[18] A. Schriver, A combinatorial algorithm minimizing submodular functions in strongly polynomial time, *J. Combinatorial Theory* **B80** (2000) 346–355.